

И.В. ПРАСОЛ, канд. техн. наук, ХНУРЭ (г. Харьков),
А.С. НЕЧИПОРЕНКО, ХНУРЭ (г. Харьков)

СПОСОБ ПОВЫШЕНИЯ РАЗБОРЧИВОСТИ РЕЧИ У БОЛЬНЫХ НЕЙРОСЕНСОРНОЙ ТУГОУХОСТЬЮ

Запропоновано новий спосіб підвищення розбірливості мови у хворих на нейросенсорну тугоухість, який базується на фільтрації мовного сигналу. В результаті фільтрації відбувається відокремлення найбільш значущих областей мовного спектру, що впливають на сприйняття мови хворим. Даний спосіб дозволяє поліпшити сприйняття мови у хворих з втратою чутності у діапазоні частот вище 1 кГц.

A new method to improve phrase intelligibility of people sick neurosurgery sensory hearing loss is suggested. It is based on the filtering voice. As a result of the filtering the most significant areas of speech spectrum which affect on speech perception patients are detected. This method helps improve the perception of speech of patients with hearing loss in the range above 1 kHz.

Постановка проблемы. В настоящее время всё большее число людей страдают нарушениями слуха. Причины снижения слуха могут быть разными, это и длительное воздействие чрезмерного шума, и общее ухудшение экологии, и осложнения после перенесенных болезней, травмы, генетические нарушения. Среди патологий слуховой системы лидирует нейросенсорная тугоухость, – ослабление слуха с сохранившимся восприятием речи, обусловленное поражением звуковоспринимающего аппарата или центрального отдела слухового анализатора. Она может быть следствием поражения как нейросенсорных эпителиальных клеток спирального органа, так и подкорковых, и корковых слуховых центров [1]. Однако чаще всего нейросенсорная тугоухость обусловлена патологией рецептора и корешка преддверно-улиткового нерва. Данное заболевание не излечимо хирургическим путем, поэтому больному назначают слухопротезирование. Слуховой аппарат должен скорректировать индивидуальные особенности слуха, характерные для данного вида патологии [2]. Для этих целей существует ряд алгоритмов обработки звука, реализованных в аналоговых слуховых аппаратах. Это алгоритмы частотно зависимого усиления, амплитудной компрессии, фильтрации помех. Однако реализация сложных алгоритмов обработки речевых сигналов, адаптируемых к индивидуальному характеру потерь слуха у пациентов, возможна только в цифровых слуховых аппаратах [3]. Цифровая обработка сигналов позволяет применять высокоэффективные адаптивные алгоритмы с возможностью индивидуальной настройки параметров при неизменном аппаратном ядре.

Анализ литературы. Проведен анализ методов цифровой обработки сигналов [4]. Изучены особенности восприятия речевого сигнала, существующие методы выделения формант из его спектра [5], а также методы

обработки речевого сигнала, позволяющие повысить разборчивость сигнала как для пациентов с нормальным слухом, так и для слабослышащих пациентов с диагнозом нейросенсорная тугоухость [6]. Однако существующие методы повышения разборчивости речи не учитывают такую особенность нейросенсорной патологии как потеря слышимости в диапазоне свыше 1 кГц. Формантный анализ не использовался для создания адаптивных алгоритмов повышения словесной разборчивости.

Цель данной статьи – разработать способ повышения разборчивости речи у больных нейросенсорной тугоухостью.

Разборчивость речи у больных нейросенсорной тугоухостью. Известно, что нейросенсорная тугоухость характеризуется повышением порогов слышимости, а также ухудшением слышимости в диапазоне свыше 1 кГц, что, в свою очередь, приводит к потере восприятия у таких больных высокочастотных компонент речи. Следствием такого ограниченного восприятия является значительное ухудшение разборчивости речи. Причина этого кроется в особенности локализации звуков речевого сигнала, а именно: согласные звуки в большинстве своём находятся в области частот выше 1 кГц, а гласные – в низкочастотном диапазоне. Поскольку на разборчивость речи преимущественно влияет восприятие согласных, то оставшийся речевой спектр из-за превалирования в нём гласных звуков становится нечётким и трудным для понимания. Исходя из вышесказанного, предлагается улучшить разборчивость речевого сигнала путём увеличения чёткости за счёт удаления частей спектра.

Особенности восприятия речевого сигнала. Любой речевой сигнал состоит из простейших звуков речи, называемых фонемами [7]. Известно, что каждой фонеме соответствует своя форма голосового тракта человека, которая варьируется в соответствии с изменением положения языка, губ, зубов, в зависимости от этого меняются и частоты голосового тракта, в том числе и резонансные. Области спектральных максимумов речевого сигнала, соответствующие резонансным частотам голосового тракта, называются формантами. Форманта характеризуется частотой, шириной и амплитудой [7]. Распознавание фонем базируется на восприятии формант как наиболее информативных признаков речевого сигнала. Также следует отметить, что распознавание каждой фонемы происходит в основном по положению первых двух формант. Учитывая данные факторы, а также такую особенность больных нейросенсорной тугоухостью как снижение частотной селективности, предлагается удалить из речевого спектра области, граничащие с формантными полосами.

Фильтрация речевого сигнала. В ходе эксперимента обрабатывалось несколько различных сигналов разных дикторов. Речевой сигнал преобразовывался в цифровой код и обрабатывался на ЭВМ. Для

идентификации формантных пиков гласных был использован программный пакет «Cool Edit Pro» [8], который позволяет проводить спектральный анализ речевого сигнала. Для получения спектральной характеристики аналогового сигнала, представленного в дискретной форме было использовано быстрое преобразование Фурье (БПФ), результатом которого является представление сигнала в виде набора из n частот

$$F = \{F_1, F_2, \dots, F_n\}. \quad (1)$$

Влияние шумовых источников на речевой сигнал может привести к образованию пиков моментального спектра, не являющихся частью речевого сигнала. Такие пики называются ложными [9].

Определим пик как максимум интенсивности энергии сигнала в определенном интервале d на оси частот и выразим функцию $P(F_k, d)$ проверки максимума в интервале d следующим образом:

$$P(F_k, d) = \begin{cases} 1, & \text{при } A_k > \max_{k-d \leq j \leq k+d} A_j, k \neq j, \\ 0, & \text{при } A_k \leq \max_{k-d \leq j \leq k+d} A_j, k \neq j. \end{cases} \quad (2)$$

Тогда нахождение всех пиков сведется к нахождению частот разложения $F_i \in F$, для которых выполняется условие $P(F_i, d) = 1$. Данный способ нахождения пиков называют последовательным проходом.

Поскольку наибольшее влияние на распознавание речевого сигнала оказывают первые две форманты, осуществляем фильтрацию формантных областей первых двух пиков. На рис. 1 изображен моментальный спектр сигнала (одна выделенная форманта) до фильтрации. На рис. 2 изображен моментальный спектр после фильтрации. Фильтрация производилась по формулам:

$$F_a = F_k - \frac{F_{\max} - F_1}{2}, \quad (3)$$

$$F_b = F_{\max} + \frac{F_2 - F_{\max}}{2}, \quad (4)$$

где F_a , F_1 , F_b , F_2 – границы формантных областей подлежащих удалению; F_{\max} – пиковая частота форманты.

$$F_b = F_{\max} + \frac{F_2 - F_{\max}}{2}, \quad (4)$$

В результате фильтрации происходит выделение наиболее значимых областей речевого спектра, а также уменьшение маскировки нижних граничных частот последующих полос, оставшихся формантными полосами.

Для создания необходимых фильтров был проведен частотный анализ гласных и согласных звуков. Спектральная огибающая содержит большое число пиков отдельных частот, однако большая их часть неинформативна и находится в области выше 1 500 Гц. Основную речевую информацию несут в

себе пики с относительно большой амплитудой в диапазоне 70 Гц – 900 Гц [10]. На рис. 3 представлен спектр звука "Е", где отчётливо различимы формантные пики в данном диапазоне.

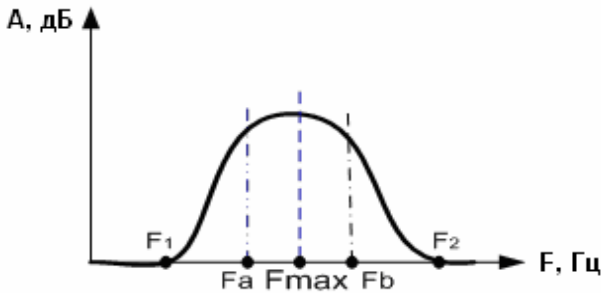


Рис. 1. Формантная линия до фильтрации

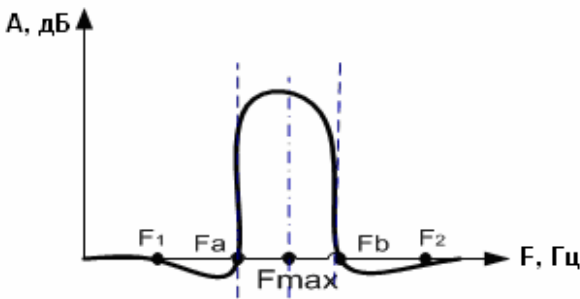


Рис. 2. Формантная линия после фильтрации

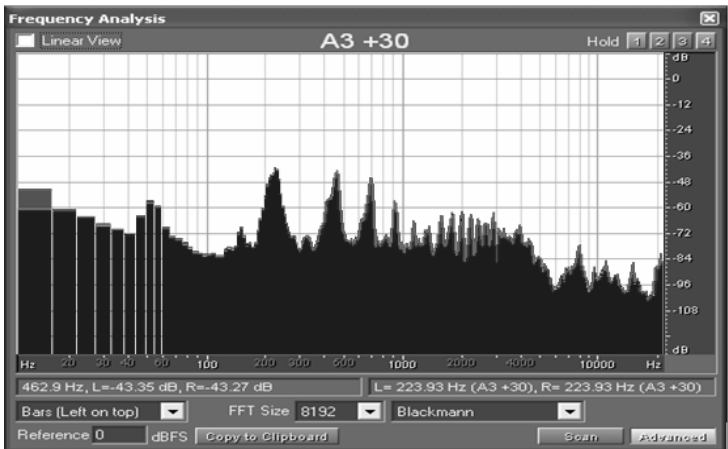


Рис. 3. Спектр звука "Е" до фильтрации

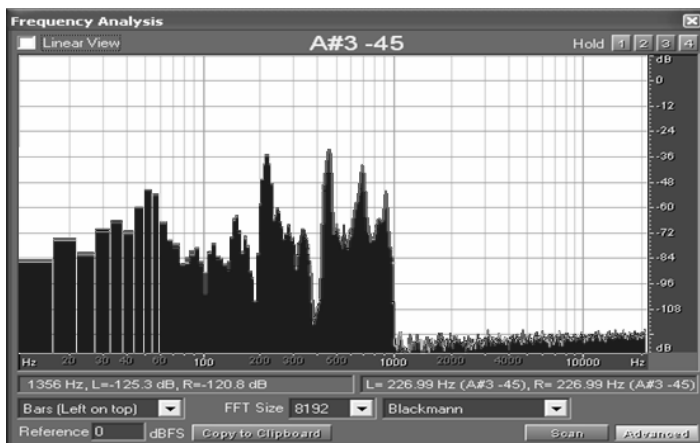


Рис. 4. Спектр звука "Е" после фильтрации

Удаление частотных полос осуществлялось с помощью фильтров, которые создавались в опции меню FFT-фильтр, согласно расчётным данным. В результате получена спектральная характеристика сигнала, представленная на рис. 4, откуда можно сделать вывод о сужении формантных областей, а также об увеличении амплитуды сигнала в области первых двух формантных пиков.

На рис. 5 – 8 представлены спектральные характеристики звуков Ж, Б до и после фильтрации.

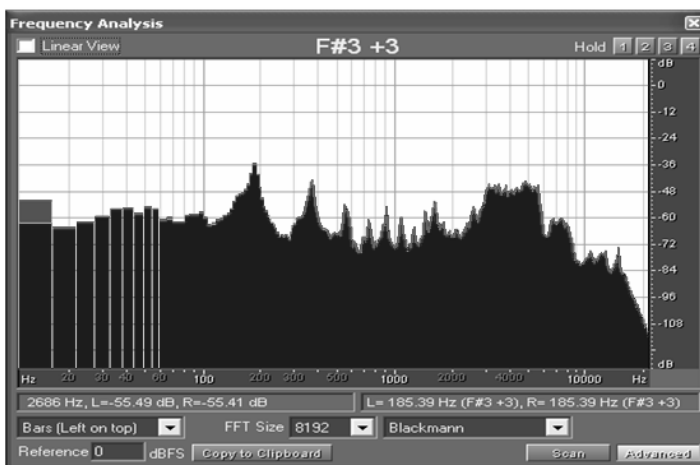


Рис. 5. Спектр звука "Ж" до фильтрации

Таким образом проводилась обработка каждого звука всей речевой последовательности. Поскольку падение кривой слышимости у больных нейросенсорной тугоухостью начинается с 1кГц, на спектр также накладывался фильтр, отсекающий частоты выше 1кГц.

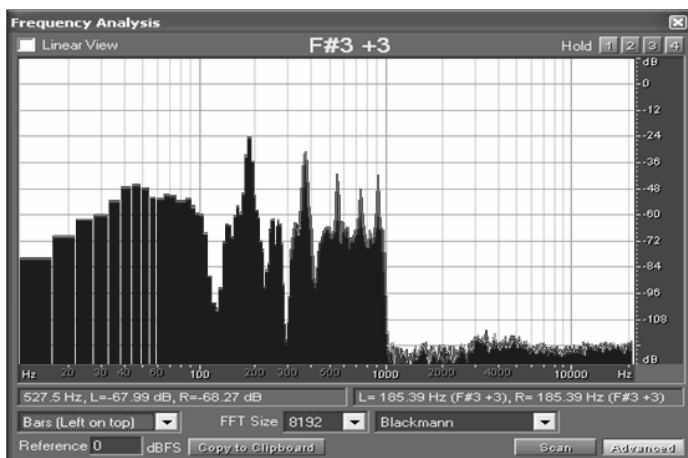


Рис. 6. Спектр звука "Ж" после фильтрации

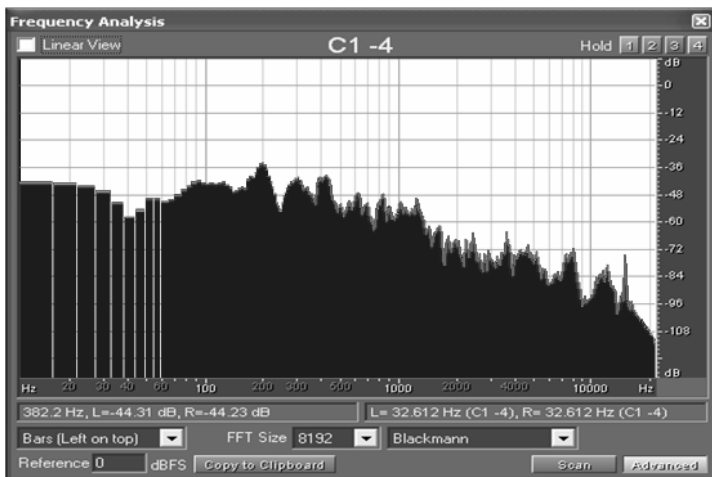


Рис. 7. Спектр звука "Б" до фильтрации

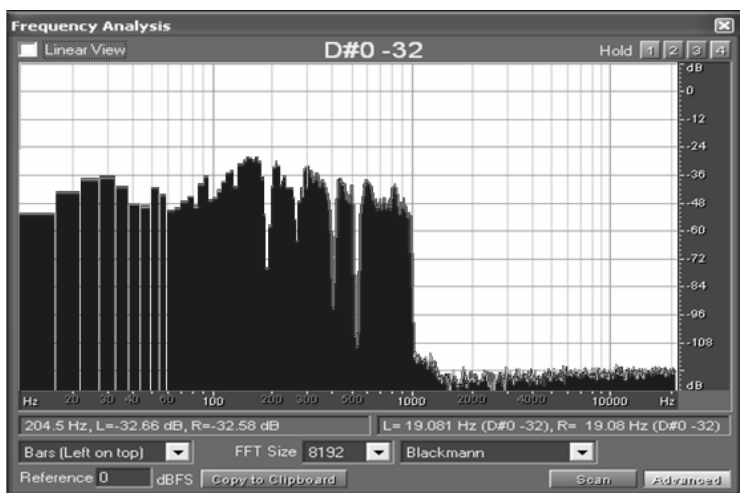


Рис. 8. Спектр звука "Б" после фильтрации

Оценка влияния предлагаемого способа обработки на восприятие речи. Для оценки влияния предлагаемого способа обработки на восприятие речевого сигнала группе больных нейросенсорной тугоухостью из 20 человек было предложено прослушать и сравнить два сигнала: исходный и прошедший предварительную обработку. Все больные имели опыт пользования слуховыми аппаратами в течение ряда лет. Следует отметить, что больные обладали невысокой исходной разборчивостью речи (менее 51%). Для оценки разборчивости были использованы речевые таблицы многосложных слов Гринберга, учитывающие фонетические особенности, характерные для русской речи [11]. Выходной сигнал не подвергался никаким дополнительным преобразованиям, кроме регулировки громкости для комфортного восприятия конкретным пациентом. Результаты исследования для 6 человек приведены в таблице.

Таблица

Результаты исследования

Испытуемые с диагнозом нейросенсорная тугоухость		Разборчивость исходного речевого сигнала (%)	Разборчивость обработанного речевого сигнала (%)
Степень тугоухости	Возраст испытуемого		
IV	50	49	54
IV	71	45	55
II	43	50	50
III	45	50	65
IV	54	41	59
III	72	48	58

Выводы. Анализируя полученные данные можно сделать вывод о повышении разборчивости речи у больных в пределах 5 – 18 %. Результаты исследования для остальных пациентов аналогичны, значения разборчивости находятся в пределах вышеуказанного диапазона. Следовательно, в результате фильтрации частотных полос спектра, граничащих с формантами, увеличивается чёткость речевого сигнала. Выделение наиболее информативных областей спектра способствует улучшению восприятия речевой последовательности, речь становится более внятной и понимаемой. Использование предлагаемого способа обработки речевого сигнала позволяет достоверно повысить разборчивость речевого сигнала у больных нейросенсорной тугоухостью. Таким образом, это открывает возможность создания адаптивного алгоритма повышения словесной разборчивости.

Список литературы: 1. *Нейман Л.В., Богомильский М.Р.* Анатомия, физиология и патология органов слуха и речи. – "Владос", 2001. – 224 с. 2. *Прасол И.В., Кобылинский А.В.* Методика оптимизации цифровых схем биомедицинских устройств // Прикладная радиоэлектроника. – 2007. – Т. 6. – № 1. – С. 51–55. 3. *Прасол И.В., Кобылинский А.В.* Алгоритмы проектирования адаптивных цифровых слуховых аппаратов / 7-я научно-техническая конференция "Медико-технические технологии на страже здоровья". Сборник научных трудов. – М.: МГТУ им. Н.Э. Баумана, 2005. – С. 54–56. 4. *Рабинер Л., Гоулд Б.* Теория и применение цифровой обработки сигналов. – М.: Мир, 1978. – 848 с. 5. *Гельфанд С.А.* Слух: Введение в психологическую и физиологическую акустику. – М.: Медицина, 1984. 6. *Петров С.М.* Полосовая фильтрация речевого сигнала – восприятие речи в норме и при нейросенсорной тугоухости // Вестник оториноларингологии. – 2000. – № 3. – С. 55–56. 7. *Рабинер Л.Р., Шафер Р.В.* Цифровая обработка речевых сигналов. – М.: Радио и связь, 1981. – 496 с. 8. <http://www.adobe.com/special/products/audition/syntrillium.html>. 9. *Сердюков В.Д.* Опознавание речевых сигналов на фоне мешающих факторов. – Тбилиси: Наука, 1987. – 142 с. 10. *Чистович Л.А., Венцов А.В., Гранстрем М.П.* Физиология речи. Восприятие речи человеком. – Л.: Наука, 1976. – 388 с. 11. *Джеймс Л. Фланаган* Анализ, синтез и восприятие речи. – М.: Связь, 1968. – 396 с.

Поступила в редакцию 20.09.2007